CMPT 374, Fall Term, 2002
Midterm Examination

Department of Computer Science
University of Saskatchewan

Friday, October 25th, 2002
Closed Book, Open Mind

22+5+10+7+22 = 66%

Overview

The total number of marks for this examination is 100, and you have 50 minutes to complete the exam. This gives you an average of 2 marks per minute. Please write all multiple choice answers on the opscan sheet provided, and answer all other answers on this booklet - if you need additional writing material raise your hand and the instructor will come by. Do not leave your seat until you are ready to hand in your exam, if you have questions raise your hand. A sheet of figures is included as the last page of this exam; you may remove this sheet at your convenience.

Part I: Multiple Choice (25 marks)

Choose the best single answer from the list of possible answers. There are no penalties for guessing.

1. The ANSI/SPARC Architecture is made up of three layers. Listed from the most abstract (end user view) to most concrete (DBMS view) they are:
    - [✓ a.] External, conceptual, internal
    - b. Internal, conceptual, external
    - c. Conceptual, external, internal
    - d. Internal, external, conceptual

2. To physically create a set of entities, attributes, and relationships in a DBMS we use a(n):
    - [✓ a.] DDL
    - b. DML
    - c. UML
    - d. ALTER

3. Which of the following is not one of Codd's eight essential functions for DBMSs:
    - a. Concurrency control
    - b. Authorization services
    - c. Transaction management
    - [✓ d.] Workflow management

4. A _______ DBMS architecture often leads to lots of network traffic while a ______ DBMS architecture tends to reduce network traffic.
    - a. Client-server, File-server
    - [✓ b.] File-server, Client-server
    - c. Multiple-server, Proxy-server
    - d. Proxy-server, Multiple-server

5. What is a relational schema?
    - a. A named relation defined by a set of attribute and domain name pairs
    - b. A set of relations each with a distinct name
    - c. An attribute, or set of attributes, that uniquely identifies a tuple within a relation
    - [✓ d.] A collection of normalized relations with distinct names

6. Can a primary key be a foreign key?
   a. Yes
   b. No
   [✓ c. No, only candidate keys can be foreign keys]

7. A trigger is most commonly used to support which kind of integrity?
   a. Entity integrity
   b. Referential integrity
   [✓ c. Enterprise constraints]

8. What integrity principle governs the following question statement: "Foreign keys must either be null or link to a candidate key in another relation"?
   a. Entity integrity
   [✓ b. Referential integrity]
   c. Enterprise constraints

9. SQL is a:
   a. First Generation Language
   b. Second Generation Language
   c. Third Generation Language
   [✓ d. Fourth Generation Language]

10. The statement "R and S are union compatible" means:
   a. R and S have the same number of tuples
   [✓ b. R and S have the same number of attributes]
   c. Either R is a subset of S or S is a subset of R
   d. R and S have the same schema

11. What is the following symbol ⟕ ?
   a. Cartesian product
   b. Natural Join
   [✓ c. Left Outer Join]
   d. Right Outer Join

12. When sorting a set of data in SQL using the ORDER BY clause, where are nulls sorted to?
   a. The top of the list
   b. The bottom of the list
   [✓ c. Either the top or the bottom of the list, depending on the DBMS]
   d. They are not sorted, they are remove
   e. They are sorted to the "N" section

13. Are you allowed to use the ORDER BY clause in a subquery?
   a. Yes
   [✓ b. No]

14. A "Fan Trap" is:
   [✓ a. Where a database model represents a relationship between two entities through another entity but that relationship is ambiguous]
   b. Where a database model implies a transitive relationships but actuality the relationship does not always exist
   c. A partially updatable view
   d. None of the above

15. Are recursive relationships allowed when creating EER diagrams?
    - [✓ a.] Yes
    - b. No

16. What is the degree of the relation shown in figure one?
    - a. One
    - b. Three
    - [✓ c.] Four
    - d. Five

17. What is the cardinality of the relation shown in figure one?
    - a. One
    - [✓ b.] Three
    - c. Four
    - d. Five

18. What normal form is the relation given in figure two in (primary key is the attribute "SerialNumber")?
    - a. Unnormalized
    - b. First Normal Form
    - c. Second Normal Form
    - [✓ d.] Third Normal Form

19. How many candidate keys are in the relation shown in figure two?
    - a. One
    - [✓ b.] Four
    - c. Five
    - d. Twenty Four

20. Given figure one, which relational algebra statement below is equivalent to this SQL statement: SELECT PricePaid FROM Figure1 WHERE CarBought = 126?
    - a. $\pi_{PricePaid}(\sigma_{CarBought}(Figure1) = 126)$
    - [✓ b.] $\pi_{PricePaid}(\sigma_{CarBought = 126}(Figure1))$
    - c. $\sigma_{PricePaid}(\pi_{CarBought = 126}(Figure1))$
    - d. $\pi_{CarBought = 126}(\pi_{PricePaid}(Figure1))$

21. Given the tables in figure one and two, what is the cardinality of the following expression FigureOne X FigureTwo?
    - a. One
    - b. Three
    - c. Six
    - [✓ d.] Nine

22. How many attributes are there when you Figure 1 $\cap$ Figure 2
    - a. Nine
    - b. Five
    - c. Four
    - [✓ d.] None, it is an invalid operation

23. Give figure three, what relational operation is diagram A referring to?
    a. Selection
    b. Projection
    c. Set Difference
    d. Union
    [✓ e.] Intersection
    f. Cartesian product
    g. Division

24. Give figure three, what relational operation is diagram B referring to?
    a. Selection
    b. Projection
    [✓ c.] Set Difference
    d. Union
    e. Intersection
    f. Cartesian product
    g. Division

25. Give figure three, what relational operation is diagram C referring to?
    a. Selection
    b. Projection
    c. Set Difference
    [✓ d.] Union
    e. Intersection
    f. Cartesian product
    g. Division

Part 2: Short Answer (25 marks)

26. What is the difference between first and second normal form? 5 marks.

The first normal form just ensures that every tuple is unique. The first normal form allows partial dependencies in the relation. The second normal for doesn't allow these partial dependencies. They must be dependent on the whole key, for 2NF.

4

27. Compare and contrast file based systems with database management systems. Include examples (with explanation) of when you would use one method over the other for data storage/retrieval. 10 marks.

The advantages of a file based system is that it is cheap, storage is small, and it allows the meta data to be defined with the application that is using it. However the disadvantages of this type of system is that it duplicates data. This leads to integrity errors on which data is correct between two files. Another disadvantage is that the data is relatively isolated from other programs. Different types of file formats also are a problem with files because ideally you want data that many different applications could use. Another problem with files is that they limit the kinds of queries you can make to them.

The advantages that DBMS's hold are that all the data is centralized so integrity isn't compromised as much. There is much less of a chance to have redundancy or conflicting data with this system, given the relations are setup right. Data is also in a fixed format so that application developers can all use the same data easily. Very powerful queries can be made using a DML. The disadvantages of a DBMS are the price (ie for software, for DBA), the storage space they consume and their complexity.

I would use a file based system for smaller applications that don't require sharing of data. I would use a DBMS for a large company that hold some data such as customer list, that will be updated frequently.

10

28. Views are an important part of the relational model. Discuss both the <u>advantages and problems of using views</u>, and the <u>support that views have in SQL</u>. Use examples if you need to to clarify your arguments. 10 marks.

Advantages:

- abstraction of data, the person viewing doesn't need to see all data values, [illegible] ✓
- [illegible] can be built on other views ✓
  the view is a direct mapping between ✓ [illegible]
  [illegible] a copy.

[illegible]

- can't easily update a view.
  - there must be at least ✓ a candidate key in the attributes provided.

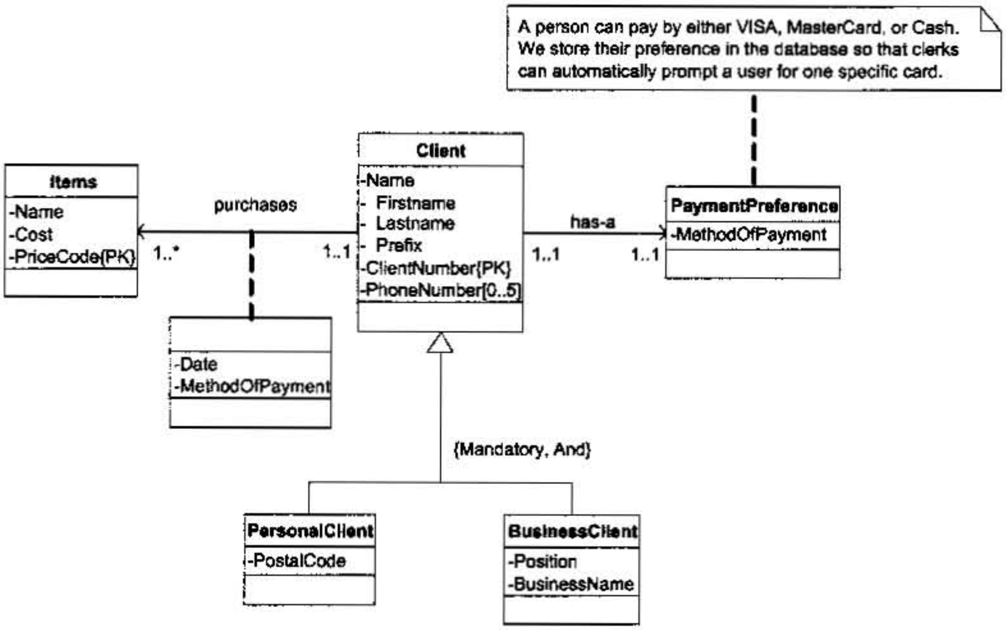
- performance?

SQL examples?

7

Part 3: Analysis and Design (50 marks)

29. Staplers: the office superstore, is looking to revamp their point of sale systems and has hired you as a database administrator and application programmer. Their Database Designer just went on vacation and has left you a copy of the high level enhanced-entity relationship model in UML notation. Your job is to take this model and transform it into a set of relations (in table form, or set form, you do not need to show SQL CREATE TABLE statements) using the techniques described in class. Provide a discussion of each technique you used.

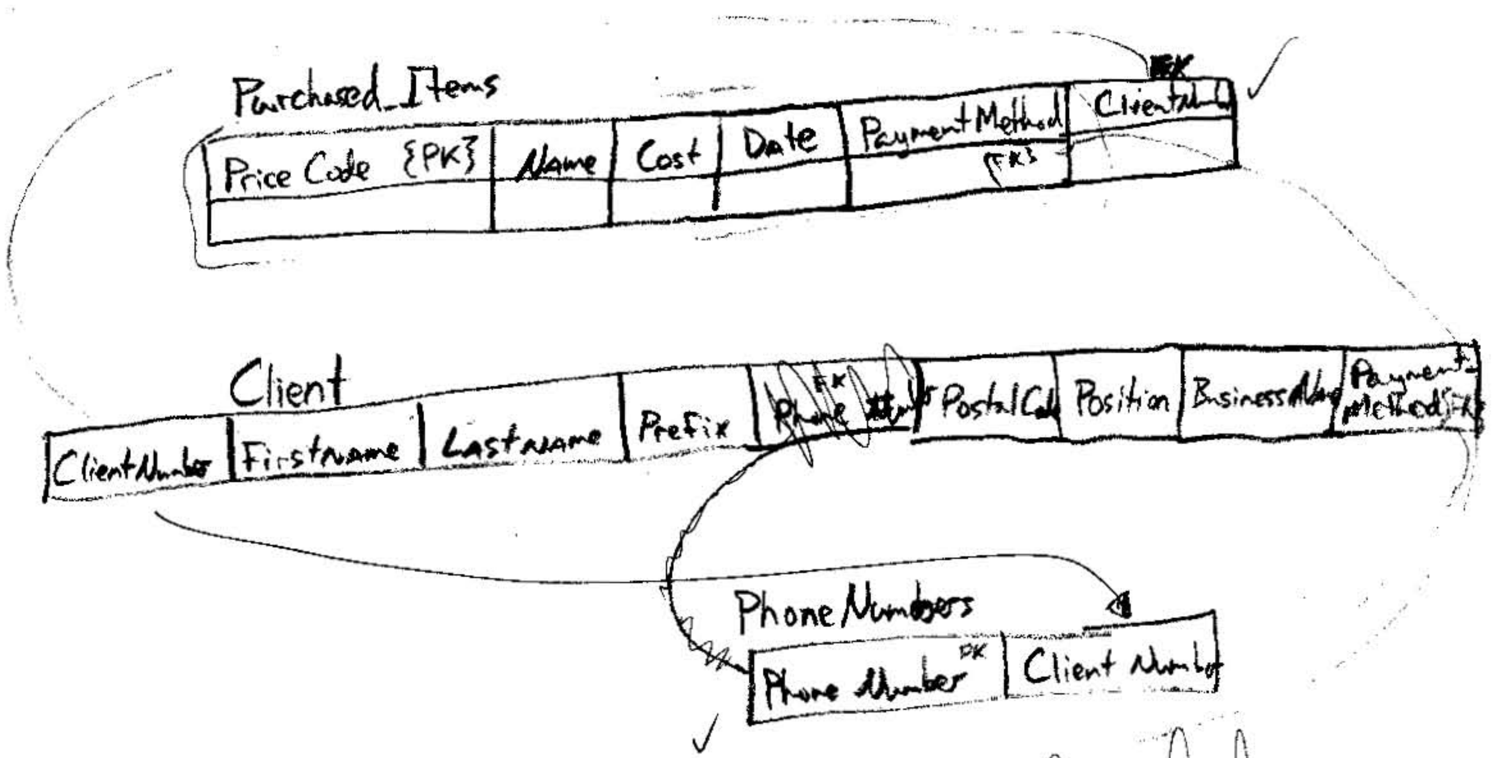


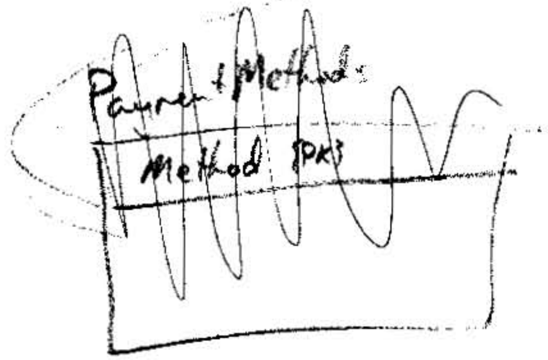
In addition to the diagram, a user who is part of the development team provides you with the following description of what might happen:

> There are two types of clients that come into Staplers: the office superstore: business clients and personal clients. All new clients are given a unique client number so the system can identify who they are. Further, we keep postal code information about personal clients and business name information about business clients for profiling reasons. Clients usually have some method of payment, which we store as a credit card name (this point of sale system is used online as well, and access to another secure system provides details of what a users credit card information is, we don't have to worry about this system).
>
> Clients come in to buy items where are all uniquely identified by a price code. A separate inventory system contains a list of all of the items in stock, but for receipt reasons we store the name of the item and the cost of the item when it is bought. We also store the date it was purchased, and how the user paid for it (again, just a credit card name or cash, for profiling reasons).

Further, provide answers (as simple SQL statements) to the following questions/statements:

a) How many purchases were done using "VISA"?
b) What is the most preferred method of payment for business customers?
c) Get a list of all of the names of customers who are in Sutherland (assume that all residents of Sutherland have a postal code that starts with "S7N").
d) It is common to send targeted fliers to people to advertise specials. Get a list of the name, address, and postal code of those personal clients who have not purchased the item "blue pen" which has a price code value of "245".





6

No explanations w. inheritance.

Need more time!

tot: 22

✓ 8

a) SELECT * FROM PurchasedItems WHERE PaymentMethod = 'VISA'

b) SELECT

0

✓

c) SELECT FirstName, LastName, Prefix FROM Client WHERE PostalCode LIKE 'S7N%'

8

d) SELECT FirstName, LastName, Prefix, address FROM (SELECT PurchasedItem p JOIN CLIENT c ON p.clientnumber = c.clientnumber) WHERE PRICECODE = '245'

We want those who have not bought a blue pen.

0

List of Figures for Multiple Choice Questions

PhoneNumber	Name	CarBought	PricePaid
306-555-9896	Billy Bob	125	$2000
306-966-4743	Cletus Slack	126	$5000
403-210-0025	Rod "Racing" Richardson	20	$10000

Figure One

Model	Colour	Year	Condition	SerialNumber
GMC Truck	Rusting Blue	1987	Excellent	125
Chrysler La Bar	Dirt Brown	1978	Good	20
Ford Ranger	Blood Red	1994	Good	126

Figure Two

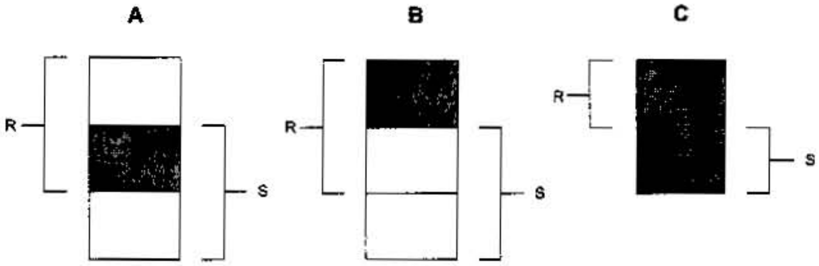


Figure Three